2020 Census Archiving Operational Assessment Report
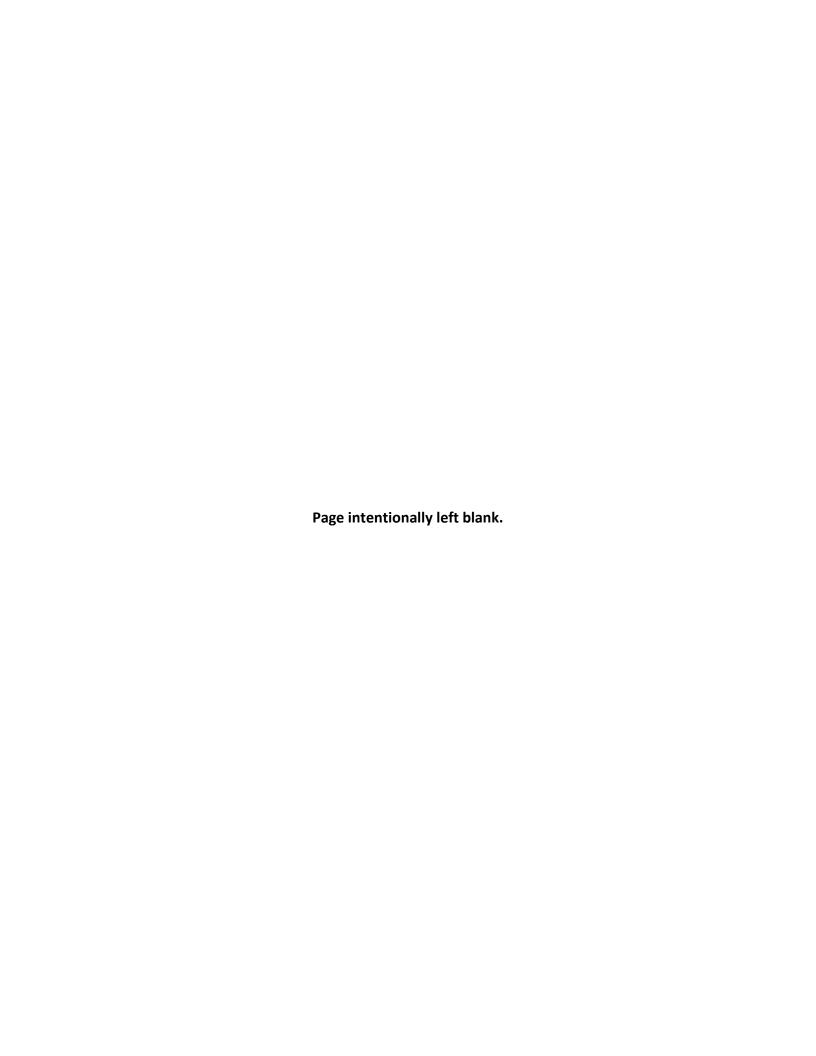
A New Design for the 21st Century

The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. (CBDRB–FY22–DSSD007–0042)

United States™ **Census** Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

United States™ **Census 2020**

Page intentionally left blank.

Table of Contents

List of Tables

Executive Summary

As of 2022, the Archiving operation (ARC) began its closeout process with an expected operational finish date in 2025. ARC's closeout will include testing the transfer of simulated data to the National Archives and Records Administration (NARA) and then ultimately transferring permanent records in production to NARA through a Simple Storage Service (S3) Bucket cloud solution. The testing of this mechanism with simulated data began in March 2022 and ended April 2022, facilitated by designated U.S. Census Bureau and NARA personnel. The process of this testing was documented in the *"2020 Census: [Census Data Lake] CDL to NARA ARC High Level Test Cases*[1]*"* document. To date, ARC has facilitated discussions with the following Census Bureau offices: Policy Coordination Office (PCO), the Records Management Office (RMO), the Office of Information Security (OIS), Decennial Architecture, and the Decennial Information Technology teams. In addition, ARC facilitated discussions with NARA's Architecture and Records Management Teams. Once testing completed, the Census Bureau and NARA began preparing for the transfer of permanent Census Bureau records in production. The parameters of this transfer can be found within the following documents created for the Census Bureau/NARA Transfer:

- Data Transfer Agreement (DTA)[2].
- Concept of Operations (ConOps)[3].
- Interface Control Document (ICD)[4].
- Interconnection Security Agreement (ISA)[5].
- Census Data Lake (CDL) Solution Architecture Diagram (SAD)[6].
- NARA's Standard Operating Procedure (SOP)[7].
- Census Archiving to NARA Production Runbook Template[8].

The remaining permanent records are expected to be sent to NARA in mid-2022 and mid-2023. These transfers include: two data transfers by the S3 Bucket cloud solution and two file transfers from the Secure File Transfer Protocol (SFTP) link (and/or DVD, depending on file size). ARC and NARA will continue to keep close communication and ensure a smooth transfer of all records deemed permanent.

---

[1] U.S Census Bureau (2022), "2020 Census: CDL to NARA ARC High Level Test Cases"
[2] U.S Census Bureau (2020), Data Transfer Agreement, September 9, 2020
[3] U.S Census Bureau (2020), Concept of Operations, June 8, 2020
[4] U.S Census Bureau (2021), Interface Control Document
[5] U.S Census Bureau (2021), Interconnection Security Agreement, September 23, 2021
[6] U.S Census Bureau (2019), CDL Solution Architecture Diagram, February 15, 2019
[7] U.S National Archives and Records Administration (2021), NARA Standard Operation Procedure, June 7, 2021
[8] U.S Census Bureau (2022), Census Archiving to NARA Production Runbook Template

The 2020 Archiving Integrated Project Team proposes the following recommendations for planning and executing the 2030 Archiving operation:

- **Begin the Archiving operation earlier in the decade for the next decennial census. In 2023, ARC should initialize and baseline the 2030 Census Master Inventory based on the 2030 Census WBS Dictionary. In 2025, ARC should initialize the 2030 Records Schedule.** Starting earlier will combat duplicative work of other operations. An early review of the records schedule will allow for a deeper dive into the previous decade's record schedule. This evaluation of the previous records schedule can help ARC draft updates of the 2030 artifacts (e.g., new products, combined artifacts, products no longer produced).

- **Start discussions about 2030 Census data products in the "7" year.** Doing so will allow for an earlier determination of where the decennial census records will be stored, how long they will be stored, the estimated size of the records, the estimated cost of storing and transferring the records to NARA, and what new records need to be archived for different purposes (i.e., NARA, research, litigation).

- **Receive approval from NARA on the 2030 Records Schedule at the same time as the Census Bureau delivers the decennial census questions to Congress (i.e., approximately in the "8" year).** Establishing an approved 2030 Records Schedule sooner than occurred for the 2020 Census will allow integrated project teams for early census operations a chance to adjust their records, if needed, and be a part of the discussion on what should be archived. These conversations will permit ARC to be proactive in establishing requirements for ARC central repositories and defining how tight controls need to be for accessing census products. Restricted records will need tighter controls for accessing and transferring data to both the central repositories and NARA.

- **Store decennial census records in one central location**. A single storage repository will simplify the transfer of records to NARA, as well as facilitate the process analysts must follow to access these records for research purposes.

- **Determine the point when archiving begins from an architectural and/or system flow standpoint.** 2020 ARC business requirements were duplicative of work being conducted in other operations. A precedent should be made that states that the Archiving operation begins when the first record is delivered to the final ARC repository. ARC recommends the operations ends after the transfer of the final 2030 ARC Records Schedule artifact is delivered to NARA.

- **Break up the "Conduct Archiving Operation" activity in the Integrated Master Schedule (IMS) into multiple separate activities**. Doing so will show a clearer path of where each record will be delivered and stored.

1. Introduction

The purpose of this document is to report on ARC's progress for the 2020 Census as well as provide further detail on the process of archiving 2020 Census data. This document was created to be used as a resource for the planning of the 2030 Census Archiving operation. The ARC, as well as its stakeholders, worked tirelessly to ensure proper storage of all 2020 Census records (i.e., records required for research, litigation, or later data releases that will occur at NARA in 72 years).

1.1 Archiving Description

The Archiving operation (ARC) performs the following functions:

- Creates the 2020 Census Records Schedule, which is a listing of all 2020 Census records, that could be delivered to the National Archives and Records Administration (NARA). These records are selected based on a list of files executed between NARA and the U.S. Census Bureau. Based on the official determination made by the Chief Archivist of the United States, records will either be permanently retained at NARA or temporarily retained at the Census Bureau until they are destroyed following established document management policy and guidelines.
- Coordinates storage of files and data and provides records deemed permanent as the official data of the 2020 Census, including sending files to NARA that contain the individual responses to the 2020 Census.
- Provides similar files to the National Processing Center (NPC) to use as source materials to conduct the Age Search Service.
- Stores data to cover in-house needs. These include data and artifacts for research and future planning purposes as well as hiring documentation that may be needed for legal purposes.

Because of the extended timeframe needed for archiving 2020 Census records and transferring them to NARA, the ARC assessment will only cover a portion of the program such as the planning, preparation, and some deliveries for ARC.

1.2 Operational Changes Resulting from COVID-19

There were no major delays for ARC, however, there were a few records within the Records Schedule that were delayed and delivered roughly two years late because of COVID changes. ARC was sure to keep NARA well-informed of all updates and changes.

1.3 Schedule

A subset of key activities/milestones for the Archiving operation from the final baselined version of the 2020 Census Integrated Master Schedule appear below in Table 1.

Table 1. Key Activities/Milestones from the Archiving Operation

Activity or Milestone Name	Planned Start	Actual Start	Planned Finish	Actual Finish
Start Archiving Project	11/03/2016	11/03/2016		
Establish Initial Baseline Review (IBR) for ARC Operational Delivery	12/21/2018	12/21/2018		
Deliver 2020 Census Records Schedule to NARA for approval			04/03/2019	04/03/2019
Finalize 2020 Census Records Schedule with NARA	04/04/2019	04/04/2019	12/09/2019	12/02/2019
Receive NARA Approval of Comprehensive 2020 Census Records Schedule	12/10/2019	12/10/2019		
Conduct Testing Readiness Review (TRR) for ARC Operational Delivery	02/14/2020	02/14/2020		
Conduct Production Readiness Review (PRR) for ARC Operational Delivery	05/08/2020	05/08/2020		
Conduct Operational Readiness Review (ORR) for ARC	07/02/2020	07/02/2020		
Conduct Archiving Operation	04/05/2019	04/16/2019	03/28/2025	
Systems Support Archiving Operation	08/03/2020	07/30/2020	09/09/2022	
Archiving Operation Finishes			03/28/2025	

Source: U.S. Census Bureau, 2020 Census, Integrated Master Schedule

2. Background

2.1 Archiving to NARA and the 2020 Census Records Schedule

The first census of the United States occurred in 1790, and a census of the population has occurred every 10 years thereafter. NARA has obtained and retained all census records for every census beginning in 1790 with the exception of 1890, for which most records were destroyed in a 1921 fire. The 1952 Exchange of Correspondence between the Archivist and the Director of the Census Bureau[9]—appended to Title 44 USC, Chapter 21.2108—outlines NARA use, custody, and maintenance of census questionnaires and microfilm, in particular, the information disclosure to the public, 72 years after the enumeration date of the census for research purposes. By law, 2020 Census results are archived and released to the public 72 years after the census. 2020 Census data records will be released by NARA in 2092.

Operational innovations for the 2020 Census include participation in cloud implementation as a solution for storing and transferring electronic records for archiving. Physical images of the 2020 Census responses were deemed temporary this decade per the 2020 Census Records

---

[9] U.S National Archives and Records Administration (1952), 1952 Exchange of Correspondence

Schedule and were not required to be transferred to the National Archives. The bulk of records deemed permanent this decade are digital records in an effort to decrease physical storage of records at NARA facilities.

2.2 Archiving for Age Search at the NPC

ARC is responsible for ensuring that the proper records are transferred to NPC for the Age Search Service. In order to conduct the Age Search operation, ARC coordinates the transfer of geographic and electronic response data as well as data capture images to NPC systems. The Census Bureau is legally authorized to provide an Age Search Service to the public. The Census Bureau has the ability to search confidential records from the federal population censuses of 1910 to 2020 and issue an official transcript of the results for a fee.

Individuals can use these transcripts, which may contain information on a person's age, sex, race, state or country of birth, and relationship to the householder, as evidence to qualify for Social Security and other retirement benefits, in making passport applications, to prove relationship in settling estates, in genealogy research, etc., or to satisfy other situations where a birth or other certificate may be needed but is not available.
For more information on the Age Search Service, click here:
<https://www.census.gov/topics/population/genealogy/agesearch.html>.

2.3 Archiving for Research and Reference

The Census Bureau is a research institution and will use some of the 2020 Census records for research and planning future censuses. Most data records are stored in central repositories and accessed through the Data Management System (DMS). Since the majority of Census Bureau records are sensitive, controlled access is required to retrieve any and all data within the central repository. Users who need to access the records must submit a request using the DMS. This request is reviewed by various levels of management and authority within the Census Bureau. Upon approval, the access privilege is granted to the user by the DMS and directed to the central repository. Other records/files are stored within the DMD Portal, which has access controls that are granted through a software license management system, REMEDY. The Decennial Management Division (DMD) Portal is accessible by all of the Decennial Census Management Division (DCMD), however, there are select folders that are restricted where only approved users have been granted access.

2.4 Overview of Archiving Definitions (Terminology)

An "archive" is a permanent repository for data and other records that have been determined to be of special importance. NARA is the federal government's record keeper. Of all documents and materials created in the course of business conducted by the United States federal

government, only 1 to 3 percent are so important for legal or historical reasons that they are kept by NARA forever. Those valuable records are preserved and are later available for use and reference by the public.

Certain Census Bureau records are determined by NARA as having sufficient historical or other archival value to warrant continued preservation by the federal government beyond the time they are needed for administrative, legal, or fiscal purposes. These records are permanently retained by NARA. Included in this group are records pertaining to the decennial census.

2.4.1   Records

The definition of "records" includes all recorded information, regardless of form or characteristics. These materials are produced by the Census Bureau pursuant to federal law or in connection with the transaction of public business. The Census Bureau preserves records as evidence of its organization, functions, policies, decisions, procedures, operations, or activities. [10]

2.4.2   Records Management

The planning, controlling, directing, organizing, training, promoting, and other managerial activities related to the creation, maintenance and use, and disposition of records, carried out in such a way as to achieve adequate and proper documentation of federal policies and transactions and effective and economical management of agency operations.[11]

2.4.3   Types of Records

Permanent Records

Permanent records are federal records that have been determined by NARA to have sufficient value to warrant their preservation in the National Archives even while they remain in agency custody. Permanent records are those for which the disposition is permanent as documented on a records schedule in the Electronic Records Archives (ERA) or standard form SF-115, Request for Records Disposition Authority, approved by NARA on or after May 14, 1973. The term also applies to all records accessioned by NARA into the National Archives of the United States.[12]

Temporary Records

---

[10] U.S National Archives and Records Administration (2017), NARA Records Management Key Terms and Acronyms
[11] U.S National Archives and Records Administration (2017), NARA Records Management Key Terms and Acronyms
[12] U.S National Archives and Records Administration (2020), Agency Permanent Records

Temporary Records are approved by NARA for disposal after a specified retention period. Records with a temporary disposition eventually will be destroyed or deleted when all relevant business needs have expired.[13]

2.4.4   Records Schedules

Records schedules provide retention and disposition instructions for Census Bureau records. NARA requires the Census Bureau to save all records identified as "permanent," including records stored in electronic media. The records schedules provide guidance for both types of records (permanent and temporary) and for the storage or disposition of each type.

The two types of records schedules are as follows:
- General Records Schedule: These schedules, published by NARA, govern the retention and disposition of administrative records common to all federal agencies.
- Census Bureau Records Schedules: These schedules govern the retention and disposition of program records.

2.5   ARC Overview

ARC is necessary to support legislative mandates. ARC supports the records management process for the 2020 Census and is responsible for producing the Census Bureau Records Schedule relating to the 2020 Census. The Records Schedule for the 2020 Census is a listing of all artifacts produced during the 2020 Census by all operations and is intended to encompass final records used to capture, process, and tabulate respondent data as well as the records used to collect and update addresses and maps.

The resulting Records Schedule has been provided to NARA, and the Chief Archivist determined whether the records are "temporary" or "permanent." Records with a temporary disposition that are in scope for ARC are destroyed by the Census Bureau once they are no longer needed to support the 2020 Census.

Response records for the 2020 Census with a permanent disposition are delivered to NARA for permanent retention, where they are preserved and stored for 72 years before they are released to the public. According to the "72-Year Rule"[14] (92 Stat. 915; Public Law 95-416), after the 2020 Census collects information from individuals, the federal government will not release

---

[13] U.S National Archives and Records Administration (2017), NARA Records Management Key Terms and Acronyms
[14] U.S. Census Bureau (2018), "The 72-Year Rule," March 02, 2018.

personally identifiable information to another individual or agency until 72 years after it was collected.

The Census Bureau will provide individual responses to the 2020 Census (including names and addresses) to NARA. The specific format, media, and timing for the delivery is negotiated between the Census Bureau and NARA. The primary use of this information is for genealogical searches (to be released no sooner than 72 years after Census Day). Permanent records were retained as defined by the retention policy. In addition, requirements were developed for retention of records needed for Census Bureau research and legal purposes. 2020 Census data have been planned for delivery to NARA in electronic format. Specific NARA staff have been trained and are dedicated to working with census records. These staff members have been granted Special Sworn Status to handle census records after demonstrating that they have suitable background clearance and have taken the Title 13 and Title 26 awareness training.

2.6   Previous Research and Literature Review

Last decade, ARC sent final census records that were in the 2010 Census Records Schedule to NARA. This involved creating the 2010 Census Records Schedule, the 2010 Records Transfer Plan, and the coordination of delivery of these records to NARA. Records were physically transferred to NARA by delivering servers.

The 2010 Archiving operation comprised of the following activities:

- Develop the 2010 Census Records Schedule.
- Approval and signature of the Records Schedule by the Chief Archivist at NARA.
- Records transfer planning.
- Coordination and monitoring of final records disposition.
- Development of a program-wide records management plan.

3.   Methodology

All 2020 Census Operational Assessments share a similar methodology. In general, they provide details about the implementation of individual operations and processes (including final volumes, rates, and costs) by presenting data from production systems, files, and activity reports, in addition to information collected from lessons learned and debriefings sessions. These important measures are key ingredients to defining successful completion of the 2020 Census operations and processes. Typical categories of success measures are as follows:

- **Process Measures** that indicate how well the process works, typically including measures related to completion dates, rates, and productivity rates.
- **Cost Measures** that drive the cost of the operation and comparisons of actual costs to planned budgets. Costs can include workload as well as different types of resource costs.
- **Quality Measures** of operational results, typically including things such as rework rates, error rates, and coverage rates.

In addition to planning and managing the implementation of its operation, each Integrated Project Team (IPT) had the responsibility of determining the assessment questions for its operation. In consultation with the Decennial Research Objectives and Methods (DROM) Working Group, each IPT developed assessment questions tailored to the uniqueness of its operation that would yield the most useful information to those planning similar operations in the future. Assessment questions provide the framework for the Results Section appearing in each operational assessment report.

The sections that follow present the assessment questions for this operation and describe the sources of information used to answer them.

1. **Process Measures** that indicate how well the process works, typically including measures related to completion dates, rates, and productivity rates.
    a. Timeframe for Records Schedule development.
    b. Milestone dates for completion of the operation.
    c. Review and record keeping of the Archiving Master Inventory.
    d. Length and scope of functionality testing.
    e. Length and scope of the production NARA data transfer.
2. **Cost Measures** that drive the cost of the operation and comparisons of actual costs to planned budgets. Costs can include workload as well as different types of resource costs.
    a. Census Image Retrieval Application (CIRA) costs for storage of images for Age Search.
    b. Age Search resource costs.
    c. ARC resource costs.
    d. NARA Transfer Costs – S3 Bucket transfer and data storage costs.
3. **Quality Measures** of operational results, typically including things such as rework rates, error rates, and coverage rates.
    a. Finalization of central repositories.
    b. Finalization and identification of Records Schedule artifacts.

    c.   Completion of Test Cases.

    d.   Completion and validation of data by NARA after transfer (expected finish four to six months after data transfer).

3.1   Assessment Questions

ARC assessment questions per the 2020 Census ARC Study Plan:

1. Has the intermediary 2020 Census Records Schedule been developed identifying all 2020 Census products that will receive a retention disposition from NARA?
2. Are all records centrally located in designated repositories and scheduled for transfer by preparing, receiving, and flagging each 2020 Census record?
3. Has the final disposition been received for all 2020 Census records?
4. What method will be used to transfer files between the Census Bureau and NARA?
5. Have all data been received by the NPC to conduct Age Search?
6. Have personnel payroll data been maintained for reference and internal research files stored for future use and in house needs?

3.2   Data Sources and Calculations: Production Systems/Reports

The files listed below were all key in planning and execution of the ARC:

Table 2. Data Sources for ARC

Data File/Report	Source	Purpose
2020 Census Records Schedule	ARC	Dispositions for 2020 Census records.
2020 Integrated Master Schedule (IMS)	ARC	ARC Schedule.
2020 Census 25. Archiving Operation Business Process Model (BPM)	ARC	ARC High Level Workflow Diagram.
Archiving Master Inventory	ARC	Internal list of 2020 Census products and storage locations for records.
2020 Census Business and Nonfunctional Requirements	ARC	List of ARC Requirements.
Census/NARA Test Cases	ARC	Detailed task descriptions and dates of the test transfer to NARA.

Data File/Report	Source	Purpose
ARC Production Runbook	ARC	Detailed task descriptions, statuses, and actual dates of the productions transfer to NARA.
Data Transfer Agreement (DTA)	ARC/CIO	Management agreement for the development, management, operation, and security of an electronic file transfer for 2020 Census records between the Census Bureau and NARA.
Concept of Operations (ConOps)	ARC/CIO	A formal document that describes how an asset, system, or capability will be used and supported.
Interface Control Document (ICD)	CDL	This document focuses on the interface between CDL and NARA.
Interconnection Security Agreement (ISA)	OIS	The Interconnection Security Agreement between U.S. Census Bureau's Census Data Lake (CDL) and NARA's Electronic Records Archives (ERA) 2.0.
CDL Solution Architecture Diagram (SAD)	CDL	Describes the technical architecture of the entire system and how each component works together to provide a solution.
NARA's Standard Operating Procedure (SOP)	NARA	This document outlines a set of Standard Operating Procedures (SOP) for how NARA intends to secure and manage 2020 Census data in the ERA 2.0 Environment in AWS (Amazon Web Service).

Source: U.S Census Bureau and National Archives and Records Administration

3.3 Lessons Learned

ARC began documenting Lessons Learned early on in the operation based on modifications from what was completed in 2010 (i.e., more schedule activities, more requirements, creating the 2020 Census Records Schedule earlier in the decade). All of these changes allowed ARC to better prepare 2020 Census operations for what NARA requires and to have NARA consider these requirements while ARC was still in the planning phase. The ARC's Lessons Learned was finalized by September 2021. All ARC members as well as ARC's key stakeholders reviewed the Lessons Learned document and received opportunities to provide input. There were many recommendations to move forward with the blueprint that the 2020 ARC team created for 2030 ARC. The ARC's Lessons Learned is attached in Appendix B.

4. Limitations

Because of the operations' long duration, the ARC assessment is limited to include only what has been completed up until 2022 and the plans for the operational closeout activities after 2022. ARC is not scheduled to be complete until 2025 and is expected to send the last 2020 Census artifact in 2035.

5. Results

ARC coordinates the transfer and storage of data created from the 2020 Census operations. The results for ARC are structured around those core goals.

5.1 Assessment Questions and Answers

ARC included the below assessment questions within the ARC Study Plan, and the updated responses based on our completed activities are below:

1. Has the Intermediary 2020 Census Records Schedule been developed identifying all 2020 Census products that will receive retention disposition from NARA?
   a. ARC conducted an annual review of records since its inception in 2016. During the annual review, ARC would meet with each operation to obtain a list of artifacts that would be expected to be kept by the National Archives or would be important for future use by the operation. The annual review consisted of a list of each 2020 Census products by operation and requested the following details: the census proposed disposition, the file format, the estimated current volume, and which census test it was included, if applicable. The 2010 Census Records Schedule was

used as a guide in this process and established a baseline of what NARA would include in the draft 2020 Census Records Schedule.

2. Are all records centrally located in designated repositories and scheduled for transfer by preparing, receiving, and flagging each 2020 Census record?

    a. A requirement was created for each of the ARC storage solutions to notify ARC that each artifact has been received and stored. Also, systems and operations include activities in their respective sections of the Integrated Master Schedule (IMS) related to archiving. The status for these activities is updated when products are sent, received and validated in order to ensure that all records were, in reality, created and properly stored. Each storage solution flagged each record with a permanent status until the official 2020 Census Records Schedule is signed by the Archivist at NARA. ARC has set up schedules and requirements this way to track where 2020 Census products are stored and where these products are located. This allows easy access when transferring records to NARA and for efficient access of records for 2030 Census planning and in-house needs. Each operation has a requirement to annually check the integrity of these files to confirm the correct files are stored and remain in a readable format.

3. Has the final disposition been received for all 2020 Census records?

    a. After the completion of the 2019 annual review of 2020 Census records, ARC deemed the intermediary 2020 Census Records Schedule ready for NARA review. The Records Schedule was then distributed to 2020 Census management to produce the Census Bureau approved 2020 Census Records Schedule. After Census Bureau approval, ARC provided the following to NARA: the Census Bureau approved 2020 Census Records Schedule and a waiver requesting early legal transfer of all permanent records before the standard 15 years. NARA points of contact formally reviewed this document, and it was signed by the Chief Archivist in December 2019. Then, ARC distributed the official 2020 Census Records Schedule to all records repositories where each disposition became final. After receiving the official Records Schedule, all ARC systems marked each record according to its final disposition.

4. What method will be used to transfer files between the Census Bureau and NARA?

    a. ARC records are being transferred by SFTP link, DVD, and AWS via the cloud. All standards have been agreed upon by both NARA and the Census Bureau and have been tested. The first data transfer by S3 Bucket containing 2020 Census data between NARA and the Census Bureau will occur in June 2022.

5. Have all data been received by NPC to conduct Age Search?

    a. The Age Search operation is using 2020 Census data. ARC confirmed that as of February 2022, all data that Age Search needed is stored within CIRA.

6. Have personnel payroll data been maintained for reference and internal research files stored for future use and in-house needs?
   a. Decennial personnel payroll data such as payroll, personnel, hiring, recruiting, retention, and internal research files are kept for in-house needs. Personnel data are stored in the following locations to address potential litigation hold:
      o Decennial Applicant, Personnel, and Payroll Systems (DAPPS).
      o Census Hiring and Employment Check System (CHEC).
      o Enterprise Census and Survey Enabling platform – Field Operational Control System (ECaSE:FOCS).
      o Enterprise Census and Survey Enabling platform – Field Operational Control System Database (ECaSE-Fld OCS: FOCS-DB).
      o Census Human Resource Information System (CHRIS).
      o Census Data Lake – Enterprise Data Lake (CDL:EDL).
   b. Files required for internal research or future census planning are stored in the following locations:
      o Decennial Statistical Studies Division (DSSD) servers.
      o Census Data Lake – Enterprise Data Lake (CDL:EDL).
   c. The Archiving products list was used to inform the internal Data Management System (DMS)—a tool to track access to datasets within the Census Bureau—and will be stored for posterity on the 2020 DMD Portal Shared Drive location within DCMD.

5.2 Validation of 2020 Census Records Listed in the Master Inventory

The tables below show the status of ARC records as of May 18, 2022:

Table 3. Status of the 2020 Census Records That Were Deemed Permanent

Number of NARA Records	
Validation	
Complete	185
In Progress	14
Not Started	
2022	4
2023	8
2024*	16
2035	1

Source: U.S. Census Bureau, 2020 Census, ARC Master Inventory
*This number is subject to change once the Data Products and Dissemination (DPD) products list is finalized

Table 4. Status of the 2020 Census Records That Are/Will Be Stored for Research

	Number of Research Records
Validation	
Complete	736
In Progress	30
Not Yet Scheduled *	22
Not Started	
2022	66
2023	15
2035	1

Source: U.S. Census Bureau, 2020 Census, ARC Master Inventory

This number is subject to change once the DPD products list is finalized

6. Conclusions and Recommendations

Even though the Archiving operation is completing an assessment, there are still a few years of work before the official closeout of the operation, according to the IMS. Although much progress has been made to ensure the proper storage of all records for the 2020 Census, there is still plenty to be done. In 2022, ARC facilitated the first of two rounds of permanent electronic data to NARA through an S3 bucket transfer. In 2023, ARC will facilitate a second data transfer to NARA for remaining permanent data products through the same S3 bucket process. There will also be more transfers using SFTP and/or DVD for remaining files/documents deemed permanent. ARC will archive all 2020 Census assessments, which are scheduled to be complete in 2024 and will facilitate its final operational closeout in 2025. Now that all transfer and storage mechanisms have been tested and used in production, all work moving forward is expected to be routine. However, if new information arises. we will add it to our lessons learned documentation.

6.1 Conclusions

All records stored at the Census Bureau will be securely stored according to Title 13 policies. Records transferred to NARA by the Census Bureau will continue to be transferred under Department of Commerce (DOC) approved protocols and methods. In addition, all permanent records will be stored and accessed at NARA using Title 13 approved policies per the interagency agreement. This plan for secure transfer and storage has been documented in a DTA that was signed by the Census Bureau and NARA.

As the official response dataset that will be transferred to NARA, the Individual Census Records File (ICRF) has been created to exclude any administrative records or Domestic Violence Shelter (DVS) data. It contains response records (self-response as well as field data collection in both

nonresponse and coverage improvement operations). This output file for NARA is based on the Census Unedited File (CUF) and received by 2020 Census staff post-production. These files do not include the application of edits or imputations introduced in later post-processing or the imputations introduced by the use of administrative records (occupied, vacant, delete, and other characteristics) and are covered under the Freedom of Information Act (FOIA) (B) (3)[15] and Title 13[16] ("Census Confidential").

The Age Search operation confirmed that all data required to conduct the Age Search Service with 2020 Census data were received as of February 2022. Now, the Census Bureau is able to search the confidential records from the federal population censuses of 1910 to 2020 and issue an official transcript of the results, when requested.

All records requested for research and litigation are still being tracked to completion and stored according to the ARC Master Inventory.

6.2 Recommendations

The 2020 Archiving operation has the following recommendations to explore for planning and executing the 2030 Archiving operation. This will allow 2030 ARC to become a more streamlined and efficient operation:

1. Determine the point in time where Archiving begins from an architectural and/or system flow standpoint. 2020 ARC business requirements were duplicative of work being conducted in other operations. A precedent should be made that states that the Archiving operation begins when the first record is delivered to the final ARC repository. ARC recommends the operations end after the transfer of the final 2030 ARC Records Schedule artifact is delivered to NARA.
2. ARC recommends determining where records will be stored and the associated cost, taking into consideration restricted (e.g., Title 13) and nonrestricted (e.g., program management documents) data and files earlier in the decade. ARC assumes that the Enterprise Storage Solution (CDL/EDL) will be the ARC repository.
3. Early decade (2023), ARC should initialize and baseline the 2030 Census Master Inventory based on the 2030 Census WBS Dictionary. Middecade (2025) ARC will initialize the 2030 Records Schedule starting with evaluating the 2020 decennial census

---

[15] Department of Justice (2004) "FOIA Guide, 2004 Edition: Exemption 3," May 2004
[16] Government Publishing Office (1954) "Title 13 – Census," August, 31, 1954

Records Schedule and review each record to determine which artifacts are no longer produced and what new records need to be archived for NARA/research/litigation.

4. The Archiving Team should evaluate the "Conduct Operation" based on the following 2020 Census milestones:
   a. Records Schedule artifacts to the Central Artifact Repository.
   b. Records Schedule data products to the Central Data Repository.
   c. Transfer of records to NARA.

5. Complete the Census Bureau-approved 2030 Records Schedule at the same time the Census Bureau submits the decennial census questions to congress (approximately in the "8" year) to ensure a timely delivery of NARA disposition and address public comment.

7. Review / Approval Table

The individuals or groups that appear in the table below have reviewed and approved this operational assessment report.

Role	Approval Date
Decennial Census Management Division (DCMD) ADC for Archiving operation	05/20/2022
Decennial Research Objectives and Methods (DROM) Working Group	06/15/2022
Decennial Communications Coordination Office (DCCO)	mm/dd/yyyy

8. Document Revision and Version Control History

The table below includes entries for each major version of this operational assessment report along with a brief description of the version and/or any changes made to the preceding version.

Version/Editor	Date	Version Description/Revisions
1.0/Keisha Mercurius	05/20/2022	Final Draft
1.0/Keisha Mercurius	06/17/2022	Final Draft with Revisions resulting from DROM comments

9. References

Government Publishing Office (1954) "Title 13 – Census," August, 31, 1954

Department of Justice (2004) "FOIA Guide, 2004 Edition: Exemption 3," May 2004

U.S. Census Bureau (2018), "The 72-Year Rule," March 2, 2018

U.S. National Archives and Records Administration (2020), Agency Permanent Records

U.S. National Archives and Records Administration (2017), NARA Records Management Key Terms and Acronyms

U.S. National Archives and Records Administration (2017), NARA Records Management Key Terms and Acronyms

U.S. National Archives and Records Administration (1952), 1952 Exchange of Correspondence

U.S. Census Bureau (2022), "2020 Census: CDL to NARA ARC High Level Test Cases"

U.S. Census Bureau (2020), Data Transfer Agreement, September 9, 2020

U.S. Census Bureau (2020), Concept of Operations, June 8, 2020

U.S. Census Bureau (2021), Interface Control Document

U.S. Census Bureau, Interface Control Document, September 28, 2021

U.S. Census Bureau (2021), Interconnection Security Agreement, September 23, 2021

U.S. Census Bureau (2019), CDL Solution Architecture Diagram, February 15, 2019

U.S. National Archives and Records Administration (2021), NARA Standard Operation Procedure, June 7, 2021

U.S Census Bureau (2022), Census Archiving to NARA Production Runbook Template

Appendix **A: Glossary of Acronyms** **Acronym**	**Definition**
ARC	Archiving operation
AWS	Amazon Web Service
BPM	Business Process Model
CDL	Census Data Lake
CDL:EDL	Census Data Lake: Enterprise Data Lake
CHRIS	Census Human Resource Information System
CIRA	Census Image Retrieval Application
CHEC	Census Hiring and Employment Check System
ConOps	Concept of Operations
COVID-19	Coronavirus Disease 2019
CUF	Census Unedited File
DAPPS	Decennial Applicant, Personnel, and Payroll Systems
DCCO	Decennial Communications Coordination Office
DCMD	Decennial Census Management Division
DMD	Decennial Management Division
DMS	Data Management System
DOC	Department of Commerce
DPD	Data Products and Dissemination
DROM	Decennial Research Objectives and Methods Working Group
DSSD	Decennial Statistical Studies Division
DTA	Data Transfer Agreement
DVD	Digital Video Disk
DVS	Domestic Violence Shelter
ECaSE-Fld OCS: FOCS-DB	Enterprise Census and Survey Enabling platform – Field Operational Control System Database
ECaSE:FOCS	Enterprise Census and Survey Enabling platform – Field Operational Control System
ERA	Electronic Records Archives
FOIA	Freedom of Information Act
IBR	Initial Baseline Review
iCADE	Integrated Capture and Data Entry
ICD	Interface Control Document
ICRF	Individual Census Records File
IMS	Integrated Master Schedule
IPT	Integrated Project Team
ISA	Interconnection Security Agreement
NARA	National Archives and Records Administration

Appendix A: Glossary of Acronyms Acronym	Definition
NPC	National Processing Center
OD	Operational Delivery
OIS	Office of Information Security
ORR	Operational Readiness Review
PCO	Policy Coordination Office
PRR	Production Readiness Review
RMO	Records Management Office
S3	Simple Storage Service
SAD	Solution Architecture Diagram
SFTP	Secure File Transfer Protocol
SOP	Standard Operating Procedure
TRR	Test Readiness Review

Appendix B: ARC 2030 Lessons Learned Planning

Number	Operation/IPT	Title	Description (include impact)	Notes	Just Do It vs. Needs Research
1	ARC	Conduct Operation Date	The Conduct Operation for Archiving should be defined as "the first time a record in the Decennial Records Schedule is finalized" (i.e. LNG, CFD)	Archivings Conduct Operation date of 8/3/20 was provided to them and did not support the work that had already occurred. The Archiving Team added a new "Conduct Operation" task to the schedule which began in April 2019 (when the first record was finalized from CFD) and renamed the 8/3/20 date as the "Systems Readiness" date.	Just Do It
2	ARC	Twin Requirements	Rather than test duplicative Archiving capability requirements (data sent from system A to central repository), Archiving will map the capability requirements in the Archiving operation to the applicable requriements from other operations (i.e. sub-od mapping).	This along with the realized versus actual workflow diagram steps will help allow us to determine if all of the testing that needs to happen is complete and if Archiving is ready to proceed to the next step in production.	Just Do It
3	ARC	Master Inventory List	Creation of the Master Inventory List which serves as the official record index for the DMS and input to the Records Schedule.	Requirements should mirror Inventory List and vice versa. The master inventory is much simpler to read and understand compared to the Decennial Records Schedule.	Just Do It
6	ARC	Age Search	Ensure there are requirements and schedule lines associated with the development and delivery of records required to "Begin Age Search with '20x0' data"	CDL (central repository) was going to be the single system to deliver response records to CIRA (Age Search system). CDL was removed as a 'pass through' system and requirements were updated to reflect the systems of record (DRPS and DRPS:IARP) sending their files to CIRA directly. Determine if the central repository will provide the Age Search records to NPC or the System of Record.	Just do it
8	ARC	Records Schedule	Begin reviewing the next Census Records Schedule in the '6' year and Approval of the Records Schedule during the '9' year	Meet individually with each IPT lead to discuss records to be added/removed - meet yearly - increase frequency in year '8'.	Just Do it
9	ARC	Requirements	Ensure PERM data that needs to be transferred to NARA are sent to the Central Repository "Unzipped". NARA has indicated they prefer to receive data files unzipped to improve the NARA Accessionning Team's validation time/effort.	Based on 6/21/21 NARA/ARC touchpoint meeting discussion	Just Do It